

University of Groningen

## Wanneer heeft selectie in het hoger onderwijs zin? De stand van zaken anno 2015

Niessen, Anna; Meijer, Rob

*Published in:*  
Tijdschrift voor Hoger Onderwijs

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Early version, also known as pre-print

*Publication date:*  
2015

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Niessen, A., & Meijer, R. (2015). Wanneer heeft selectie in het hoger onderwijs zin? De stand van zaken anno 2015. *Tijdschrift voor Hoger Onderwijs*, 33(4), 4-19.

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

## **Wanneer heeft selectie in het hoger onderwijs zin? De stand van zaken anno 2015**

A. Susan M. Niessen

Rob R. Meijer

A. Susan M. Niessen, Rob R. Meijer, Afdeling Psychometrie en Statistiek, Faculteit Gedrags- en Maatschappijwetenschappen, Rijksuniversiteit Groningen

Correspondentie over dit artikel kan geadresseerd worden aan A. Susan M. Niessen, Afdeling Psychometrie en Statistiek, Faculteit Gedrags- en Maatschappijwetenschappen, Rijksuniversiteit Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen. Contact:

[a.s.m.niessen@rug.nl](mailto:a.s.m.niessen@rug.nl)

## Samenvatting

In het Nederlandse hoger onderwijs is het aantal opleidingen met een numerus fixus en de mogelijkheden voor het selecteren van studenten sterk toegenomen. Er wordt veel onderzoek uitgevoerd naar de effecten van selectie in het hoger onderwijs. Belangrijk zijn daarbij enerzijds de toename in rendement en de kwaliteit van studenten en anderzijds het beperken van het ten onrechte afwijzen van studenten. Studies over selectie in het onderwijs verstekken echter vaak geen informatie over bovenstaande zaken, maar rapporteren slechts verschillen tussen groepen of correlaties tussen selectiemethoden en criteria. Om uitspraken te doen over de toegevoegde waarde van selectie zijn verschillende utiliteitsmodellen beschikbaar die ook rekening houden met contextfactoren zoals de toevalskans en de selectieratio. In dit artikel worden het Taylor-Russell model en het Naylor-Shine model besproken in de context van selectie in het onderwijs en toegelicht met empirische voorbeelden. Vanuit de utiliteitsmodellen kan geconcludeerd worden dat selectie alleen van toegevoegde waarde is als de toevalskans en de selectieratio niet erg hoog zijn, zelfs bij een hoge predictieve validiteit. Kijkend naar de stand van zaken in het Nederlandse hoger onderwijs, waarin de selectieratio vaak erg hoog is, zal selectie weinig tot geen toegevoegde waarde in rendement en kwaliteit van studenten opleveren.

Added value of selection in higher education in 2015

## Abstract

In Dutch higher education the number of programs with a numerus fixus and the possibility to select students have drastically increased. Many studies are being conducted to study the effects of selection in higher education. Important topics in selective admission research are (1) increasing the probability that students can successfully complete college work within the allotted time and increasing the quality of students, and (2) limiting the number of students who are falsely rejected. However, studies about selective admission usually do not provide information about these topics, but only report differences between groups of students or correlations between test scores and criterion scores. To be able to draw conclusions about the added value of selective admission, several models are available that also take context factors like base rates and selection ratios into account. In this paper the Taylor-Russell model and the Naylor-Shine model are discussed and these models are illustrated using empirical examples. Empirical values for predictive validity, base rates and selection rates in higher education in the Netherlands are also discussed. Based on the utility models it can be concluded that selection only has added value when the base rate and the selection ratio are relatively low, even when predictive validity is high. Looking at the current situation in Dutch higher education, where selection rates are often high, selection in higher education will have little or no added value in terms of student quality and study effectiveness.

*Trefwoorden:* hoger onderwijs, selectie, utiliteit, rendement

*Key words:* higher education, selection, utility, student quality

## Inleiding

In het hoger onderwijs in Nederland is het aantal opleidingen met een numerus fixus de afgelopen tien jaren behoorlijk toegenomen (Onderwijsraad, 2014). Ongeveer een derde van alle studenten meldt zich aan voor een studie met een numerus fixus (Ministerie van Onderwijs, Cultuur en Wetenschap, 2014). Ook de mogelijkheid tot het selecteren van studenten door instellingen is toegenomen. In 2017 neemt decentrale selectie definitief de rol van de centrale loting over. Dit zou tot hogere rendementen en minder studie uitval moeten leiden, omdat selectie op inhoudelijke criteria moet zorgen voor een betere match tussen de student en de studie (Wet kwaliteit in verscheidenheid, 2013).

Gezien deze ontwikkelingen is er veel aandacht voor de effecten van selectie en wordt onderzoek uitgevoerd naar methoden en effecten van selectie in het hoger onderwijs. In sommige studies worden groepen met elkaar vergeleken op basis van studieprestaties, bijvoorbeeld een groep studenten die via decentrale selectie is toegelaten en een groep die via loting is toegelaten (zie bijvoorbeeld Schripsema, van Trigt, Borleffs & Cohen-Schotanus, 2014; Urlings-Strop, Stijnen, Themmen & Splinter, 2009; Visser, van der Maas, Freeke-Engels & Vorst, 2012). Deze groepen werden echter niet random ingedeeld. Studenten konden zelf de keuze maken om deel te nemen aan een selectieprocedure of niet. Hoewel er in deze studies gecorrigeerd werd voor een aantal relevante variabelen zoals geslacht en cijfers in het voortgezet onderwijs, kunnen andere relevante variabelen zoals motivatie een rol spelen in de keuze om wel of niet deel te nemen aan een selectieprocedure (Schripsema et al., 2014). Daarom is het moeilijk om conclusies te trekken over het effect van selectie als alle studenten straks decentraal geselecteerd moeten worden voor opleidingen met een numerus fixus.

Andere studies onderzochten het voorspellen van studiesucces door de relatie tussen prestaties tijdens een selectieprocedure en studieprestaties te onderzoeken in de vorm van correlatiecoëfficiënten en regressieanalyses (zie bijvoorbeeld van der Heijden, Hessen & Wubbels, 2012; Niessen, Meijer & Tendeiro, 2014). Hoewel deze studies belangrijke vragen beantwoorden, geven zij geen antwoord op de vraag wat de *toegevoegde* waarde is van het toepassen van selectiemethoden in termen van veranderingen in het rendement en de kwaliteit van studenten. Voor het beantwoorden van die vragen zijn verschillende utiliteitsmodellen beschikbaar. Utiliteitsmodellen voor selectieprocedures zijn ontwikkeld vanuit de Arbeids- en Organisationspsychologie in de context van personeelsselectie. Zij zijn echter toepasbaar op elke selectiesituatie, maar worden zelden toegepast in de context van selectie in het onderwijs. Deze utiliteitsmodellen gebruiken naast informatie over de predictieve validiteit van een selectie instrument (de correlatie tussen prestaties op selectietoetsen en studieprestaties) ook gegevens over de context waarbinnen die selectie plaatsvindt. Die factoren zijn van grote invloed op de utiliteit van een selectieprocedure.

In dit artikel worden twee utiliteitsmodellen uitgebreid besproken. Daarnaast worden de modellen gedemonstreerd met voorbeelden op basis van recente gegevens verkregen uit een selectieprocedure bij psychologiestudenten en studieprestaties van die studenten in het eerste jaar. Zowel de stand van zaken rond de predictieve validiteit van selectieprocedures als relevante contextfactoren in het Nederlands hoger onderwijs worden besproken en implicaties voor de utiliteit van het toepassen van selectie in het hoger onderwijs worden gegeven.

## Selectiebeslissingen

Er zijn verschillende methoden en eigenschappen die geschikte voorspellers van studiesucces zijn gebleken, zoals prestaties in het voortgezet onderwijs (van der Heijden et al., 2012), persoonlijkheidsfactoren (Richardson, Abraham & Bond, 2012) en work sample tests (Meijer &

Niessen, 2015; Visser et al., 2012). Het is echter duidelijk dat geen enkele voorspeller, of een combinatie daarvan, perfect studieprestaties kan voorspellen of daar zelfs maar in de buurt komt. Eigenschappen en kenmerken van studenten worden gemeten via feilbare tests en procedures. Als gevolg daarvan worden selectiebeslissingen altijd met een zekere onzekerheid gemaakt (Mellenbergh, 1993). Mellenbergh (1993) beschrijft in dit kader de verschillende componenten van selectiebeslissingen. De eerste component is de *beslisser*, een persoon of instituut die de beslissingen neemt. Ten tweede is er een *actie*, in dit geval het aannemen of afwijzen van een student. De actie is gebaseerd op bepaald verwacht *criteriumgedrag* (in deze context de verwachte studieprestaties). Het verwachte criteriumgedrag is weer gebaseerd op *informatie*, hier de prestaties op een selectieprocedure. Die informatie wordt gebruikt in een *beslisregel*, die voorschrijft welke actie genomen moet worden gegeven de informatie. Ten slotte heeft de beslissing een zekere *utiliteit*, de effectiviteit van de combinatie van de gekozen actie en het uiteindelijke criteriumgedrag.

### Utiliteitsmodellen

Een belangrijk component van de utiliteit van selectiebeslissingen is de hierboven besproken relatie tussen de informatie en het criteriumgedrag, zoals dat in correlatiestudies onderzocht wordt. Wat echter ook relevant is, is dat wanneer er geen selectie plaatsvindt er ook een zeker rendement zal zijn. Hoewel de grootte daarvan kan variëren, zal het rendement zonder selectie niet gelijk zijn aan nul, maar vaak veel hoger. Daarnaast zijn selectieprocedures nooit perfecte voorspellers, dus zullen er ook onsuccesvolle studenten worden aangenomen als selectieprocedures worden gebruikt en bovendien succesvolle studenten worden afgewezen. Hoeveel van deze ‘verkeerde’ beslissingen gemaakt worden hangt af van de proportie sollicitanten die worden toegelaten. Deze aspecten worden vaak niet meegenomen in onderzoek naar selectieprocedures, in ieder geval niet expliciet. Wanneer het gaat om het effect van selectieprocedures, moet ook de toegevoegde waarde van een selectieprocedure tegenover het niet gebruiken ervan onderzocht worden in termen van rendement, zoals ook Drenth (1995;1998) en Mellenbergh (1995) al eens beargumenteerd hebben binnen de context van het Nederlandse hoger onderwijs.

Er zijn verschillende utiliteitsmodellen beschikbaar om de toegevoegde waarde in rendement te onderzoeken. Twee daarvan worden hier besproken. Het eerste is het model van Taylor en Russell (1939). Binnen dit model is utiliteit gedefinieerd als de toename in succesvolle studenten na het implementeren van een selectieprocedure (zie ook Drenth, 1995;1998). Het tweede model is het model van Naylor en Shine (1965). Binnen dit model wordt utiliteit gedefinieerd als de toename in gemiddelde criteriumprestatie na het implementeren van een selectieprocedure. Andere modellen zijn die van Lawshe, Bolda, Brune en Auclair (1958) en van Brogden (1946) en Cronbach en Gleser (1965). In het model van Lawshe et al. (1958) worden alle sollicitanten opgedeeld in groepen aan de hand van selectiescores (bijvoorbeeld laag, gemiddeld en hoog). Per groep kan dan de proportie succesvolle studenten worden bepaald. Met het model van Brogden (1946) en Cronbach en Gleser (1965) kan de utiliteit in termen van financiële opbrengsten worden bepaald (zie ook Mellenbergh, 1995). Het schatten van de financiële utiliteit is in de praktijk echter lastig omdat schattingen van gemiddelde kosten en spreiding nodig zijn. Het model van Lawshe et al. (1958) is een extensie van het Taylor-Russell model, waarbij er in plaats van twee groepen (afwijzen en aannemen), meerdere groepen kunnen worden onderscheiden.

Hierna worden de utiliteitsmodellen van Taylor en Russell (1939) en Naylor en Shine (1965) uiteengezet. Ook de aannamen van deze methode en de gevolgen daarvan voor de nauwkeurigheid van de schattingen van de utiliteit van selectieprocedures worden besproken.

Vervolgens worden beide modellen toegepast op selectiedata en studieprestaties van een cohort eerstejaars psychologiestudenten.

### *Taylor-Russell model*

Het Taylor-Russell model is het bekendste en simpelste utiliteitsmodel. Binnen het Taylor-Russell model wordt het criteriumgedrag gedichotomiseerd tot de groepen 'succesvol' en 'onsuccesvol', aan de hand van een gestelde criteriumscore. Utiliteit wordt uitgedrukt in een toename in de succesratio, de proportie succesvolle studenten binnen alle geselecteerde studenten (zie Figuur 1, vergelijking 1). In Figuur 2 wordt dit schematisch weergegeven. In vak A staan de personen die ten onrechte zijn afwezen (niet toegelaten, wel succesvol), in vak B staan de personen die terecht zijn toegelaten (wel toegelaten, wel succesvol), in vak C staan de personen die terecht zijn afgewezen (niet toegelaten, niet succesvol) en in van D staan de personen die ten onrechte zijn toegelaten (wel toegelaten, niet succesvol). De toename van de succesratio hangt af van drie zaken. Voor de utiliteit is ten eerste de predictieve validiteit van de selectieprocedure van belang, uitgedrukt in een correlatie tussen het selectie instrument en het criteriumgedrag. Ten tweede wordt ook de toevalskans (Engels: base rate), het rendement zonder selectie, in beschouwing genomen. De toevalskans is de proportie succesvolle studenten wanneer er geen selectie plaatsvindt (zie Figuur 1, vergelijking 2). Ten derde houdt het model rekening met de selectieratio, de proportie te selecteren studenten uit de gehele sollicitantenpool (zie Figuur 1, vergelijking 3). De predictieve validiteit van de selectieprocedure wordt weergegeven door de breedte van het ovaal (zie Figuur 2). Als informatie beschikbaar is over de toevalskans, de selectieratio en de predictieve validiteit van de selectieprocedure, kan de bijbehorende succesratio opgezocht worden in de Taylor-Russell tabellen. De utiliteit is dan de succesratio met selectie minus de toevalskans. Er is een online calculator beschikbaar om de succesratio te berekenen (McLellan, 1996.).

(Plaats Figuur 1)

(Plaats Figuur 2)

Het punt dat Taylor en Russell (1939) met hun model maakten was dat de utiliteit van een selectieprocedure niet alleen afhankelijk is van de predictieve validiteit, maar ook van de toevalskans en de selectieratio. Wanneer de toevalskans hoog is (zonder selectie zijn er al relatief veel studenten succesvol) heeft een selectieprocedure weinig toegevoegde waarde, zelfs als de predictieve validiteit hoog is. Wanneer de selectieratio hoog is (bijna iedereen wordt toegelaten, bijna niemand wordt afgewezen), zal een selectieprocedure zelfs met een hoge predictieve validiteit ook weinig utiliteit hebben. Omgekeerd kan een selectieprocedure met een relatief geringe predictieve validiteit een aanzienlijke utiliteit opleveren, wanneer de toevalskans en de selectieratio klein zijn (Taylor & Russell, 1939).

### *Aannamen Taylor-Russell model*

Het model van Taylor en Russell (1939) is eenvoudig in het gebruik en kan nuttige informatie over de utiliteit van een selectieprocedure opleveren. Wel gaat het model uit van een aantal aannamen waaraan soms moeilijk te voldoen is. Een eerste aanname binnen het Taylor-Russell model is dat de toevalskans (het aantal succesvolle studenten zonder selectie) bekend is. Dat is voor opleidingen die al langere tijd selectieprocedures gebruiken niet het geval. De succesratio onder een bestaande selectieprocedure met een validiteit groter dan nul is niet gelijk aan de toevalskans. Aannemen dat ze wel gelijkwaardig zijn leidt tot een overschatting van de utiliteit (Smith, 1948). Smith (1948) beschrijft echter wel een procedure om in het geval van een bestaande selectieprocedure waarvan de predictieve validiteit bekend is, de utiliteit van een

(nieuwe) selectieprocedure te bepalen. Een tweede aanname is dat de validiteitscoëfficiënt gebaseerd is op gegevens van huidige studenten die aangenomen zijn met een selectieprocedure met een validiteit gelijk aan nul, of zonder selectieprocedure. De sollicitanten en huidige studenten kunnen dan als gelijkwaardig beschouwd worden. Predictieve validiteit is echter meestal gebaseerd op onderzoek waarbij de te onderzoeken selectieprocedure al in werking is. Er zullen dus ook mensen worden afgewezen op basis van de selectieprocedure, wat leidt tot range-restrictie. Range-restrictie zorgt voor een onderschatting van de predictieve validiteit. Als er sprake is van range-restrictie in de steekproef op basis waarvan de predictieve validiteit berekend wordt, zorgt dat voor een onderschatting van de predictieve validiteit en als gevolg ook tot een onderschatting van de utiliteit (Smith, 1948). Er zijn echter methoden beschikbaar om validiteitscoëfficiënten te corrigeren voor range-restrictie. Meestal wordt dan gebruik gemaakt van de geobserveerde correlatie, de spreiding in de geselecteerde groep en de spreiding in de gehele groep om de eigenlijke correlatie te schatten (zie bijvoorbeeld Sacket & Yang, 2000).

Ook gaat het Taylor-Russell model uit van bivariate normale, lineaire en homoscedastische relaties tussen de selectieprocedure en het criterium<sup>1</sup>. De dichotomisering van het criterium naar succesvol en onsuccesvol is gebaseerd op een onderliggende, normaal verdeelde variabele, waarvoor een cutoff punt is vastgesteld. Wanneer het criterium een daadwerkelijke dichotomie voorstelt (bijvoorbeeld het wel of niet stoppen met een studie), zijn de Taylor-Russell tabellen ongeschikt. Abrahams, Alf en Wolfe (1971) hebben gelijkwaardige tabellen opgesteld die in dat geval gebruikt kunnen worden. Hoewel er binnen het Taylor-Russell model wordt verondersteld dat er een normaal verdeelde variabele ten grondslag ligt aan de dichotomisering naar succesvol en onsuccesvol, worden alle succesvolle studenten binnen dit model als even succesvol beschouwd (Cascio, 1980). Er kunnen geen uitspraken gedaan worden over hoeveel succesvoller studenten worden als er sprake is van selectie, alleen hoeveel meer succesvolle studenten er zullen zijn. Cascio (1980) noemt dit als groot nadeel van het Taylor-Russell model en stelt dat dit leidt tot een onderschatting van de utiliteit van een selectieprocedure als men verwacht dat binnen de groep succesvolle studenten ook onderscheid te maken valt tussen meer en minder succesvolle studenten.

#### *Naylor-Shine model*

Het Naylor-Shine model (Naylor & Shine, 1965) vereist geen dichotomisering van het criteriumgedrag, waardoor onderscheid gemaakt kan worden in de mate van succes. Het rendement wordt uitgedrukt via de gemiddelde criteriumprestatie. In het geval van studenten kan dat bijvoorbeeld het gemiddelde cijfer zijn. De utiliteit van een selectieprocedure wordt dan uitgedrukt als de toename in gemiddelde criteriumprestatie (gemiddelde criteriumprestatie van de geselecteerde groep minus de gemiddelde criteriumprestatie van de gehele groep sollicitanten), gegeven de predictieve validiteit en de selectieratio. De utiliteit kan worden opgezocht in tabellen (Naylor & Shine, 1965), of berekend worden met een online calculator (McLellan, 1996). Het Naylor-Shine model gaat uit van een lineaire relatie tussen predictieve validiteit en utiliteit. Daarnaast geldt dat hoe lager de selectieratio, hoe hoger de utiliteit. Ook hier wordt een bivariate normale, lineaire en homoscedastische relatie tussen de selectieprocedure en het criterium aangenomen. Ook geldt de aanname dat de predictieve validiteit is bepaald op basis van huidige studenten, die zijn aangenomen zonder selectie of met een selectieprocedure met een validiteit gelijk aan nul (Cascio, 1980). Er wordt dus uitgegaan van een

---

<sup>1</sup> Een impliciete aanname is ook dat het verband tussen de selectieprocedure en het criterium gelijk is voor alle personen. Wanneer er sprake is van differentiële predictie op basis van bijvoorbeeld geslacht of sociale kenmerken, zal ook de utiliteit verschillen voor deze groepen. Dat zal in de ene groep leiden tot meer ten onrechte aangenomen en afgewezen studenten dan in de andere groep. Dit is uiteraard onwenselijk en dient vooraf onderzocht te worden.

validiteitscoëfficiënt die verkregen is zonder range-restrictie en dat de gemiddelde criteriumprestatie van de gehele sollicitantengroep bekend is. Het Naylor-Shine model kan voor twee doelen gebruikt worden. Ten eerste kan de gemiddelde toename in de criteriumprestatie gegeven een selectieprocedure met een bekende predictieve validiteit en selectieratio bepaald worden. Ook kan er een ideale cutoff score voor een selectie instrument berekend worden, gegeven een gewenste gemiddelde criteriumscore in de geselecteerde groep. Met dit model kan dan de toename in het gemiddelde cijfer van studenten met een bepaalde selectieprocedure bepaald worden. Ook kan berekend worden wat de cutoff score in de selectieprocedure moet zijn als het doel is om het gemiddelde cijfer van studenten, bijvoorbeeld, een zeven te laten zijn.

### Empirische voorbeelden

Op basis van gegevens van eerstejaarsstudenten psychologie aan de Rijksuniversiteit Groningen worden twee voorbeelden gegeven. Eerst wordt de utiliteit van een selectieprocedure voor het criterium stoppen met de studie in het eerste jaar gedemonstreerd aan de hand van het Taylor-Russell model, maar gebruikmakend van de aangepaste tabellen van Abrahams et al. (1971), aangezien stoppen met de studie een dichotoom criterium is in plaats van een dichotomisering van normaal verdeelde scores. Ook wordt het Naylor-Shine model toegepast om de mate van stijging in succes naar aanleiding van een selectieprocedure te demonstreren. Als criterium geldt dan het gemiddelde cijfer in het eerste jaar. Er worden resultaten getoond voor twee selectieratio's, één voor een strenge en één voor een milde selectie. Voor de milde selectie wordt uitgegaan van een selectieratio van .80, wat inhoudt dat 80 procent van alle sollicitanten zou worden toegelaten. Voor de strenge selectie wordt uitgegaan van een selectieratio van .30, waarbij 30 procent van de sollicitanten zou worden toegelaten.

We gebruiken gegevens van 653 studenten psychologie die voorafgaand aan de start van de studie in 2013 hebben deelgenomen aan een selectieprocedure waarin ze een toets hebben gemaakt. De studenten kregen daarvoor beschikking over twee hoofdstukken studiemateriaal van het vak *introductie in de psychologie*, met de opdracht de stof thuis te bestuderen. Vervolgens hebben zij op de toetsdag een tentamen van 40 items gemaakt over de stof. De betrouwbaarheid van de toets was  $\alpha = .81$ , met een gemiddelde score van  $M = 30$ ,  $SD = 5,2$  (zie voor een uitgebreide beschrijving van de procedure Meijer & Niessen, 2015; Niessen et al., 2014). Na afloop van de selectieprocedure zijn alle studenten toegelaten. Studenten die niet aan de studie zijn begonnen na deelname aan de selectieprocedure hebben daar zelf voor gekozen. Daarom is de predictieve validiteit van de selectietoets zonder range-restrictie bekend en is de toevalskans voor stoppen met de studie bekend, er is immers niet daadwerkelijk geselecteerd<sup>2</sup>. Daarmee voldoen onze gegevens aan de aannamen van beide modellen. De correlatie (afgerond) tussen de scores op de selectietoets en succes (niet stoppen met de studie) was  $r = .3$  (een redelijk verband) en de correlatie tussen de score op de selectietoets en het gemiddelde cijfer in het eerste jaar was  $r = .5$  (een sterk verband). De proportie studenten die in het eerste jaar niet gestopt is met de studie en dus succesvol was, was .80. Dit is de toevalskans in onze steekproef. Het gemiddelde cijfer in het eerste jaar was  $M = 6,6$  ( $SD = 1,3$ ).

#### *Utiliteit voor het criterium stoppen met de studie*

Gebruikmakend van de aangepaste Taylor-Russell tabellen van Abrahams et al. (1971), wordt duidelijk dat deze selectieprocedure met een predictieve validiteit van  $r = .3$ , bij een toevalskans van .80 en een selectieratio van .80, een succesratio van 0.85 oplevert. Bij deze milde selectie zullen er 5 procent minder studenten stoppen met de studie dan het geval zou zijn zonder

<sup>2</sup> Het is mogelijk dat de sollicitanten die er voor kozen om niet aan de studie te beginnen verschilden van studenten die wel begonnen op bijvoorbeeld sociale kenmerken. Hierover waren geen gegevens beschikbaar.



selectie. In deze situatie zou van alle sollicitanten 68 procent terecht worden toegelaten en 8 procent terecht worden afgewezen, maar ook 12 procent ten onrechte worden toegelaten (wel toegelaten maar stoppen met de studie) en 12 procent ten onrechte worden afgewezen (niet toegelaten maar waren niet gestopt met de studie). Zonder selectie wordt 80 procent terecht toegelaten en 20 procent ten onrechte toegelaten. Er wordt dan niemand, terecht of onterecht, afgewezen. Het invoeren van deze selectieprocedure met een selectieratio van .80 zou dus zorgen voor een kleiner aantal terecht toegelaten studenten, maar ook voor een kleiner aantal ten onrechte toegelaten studenten. Dit scenario is nog eens grafisch weergegeven in Figuur 3a.

Wanneer wordt uitgegaan van een strenge selectie met een selectieratio van .30 is de bijbehorende succesratio.92. Deze strenge selectie zou er voor zorgen dat er 12 procent minder studenten stoppen met hun studie ten opzichte van een situatie zonder selectie. Bij een strenge selectie zou van alle sollicitanten 28 procent terecht worden toegelaten en 18 procent terecht worden afgewezen, zou 2 procent ten onrechte worden toegelaten en 52 procent ten onrechte worden afgewezen. In het geval van strenge selectie neemt het rendement dus sterker toe, maar neemt het aantal studenten dat ten onrechte wordt afgewezen ook sterk toe ten opzichte van een milde selectie. In Figuur 3 wordt dit nog eens grafisch weergegeven.

(Plaats Figuur 3)

#### *Utiliteit voor het criterium gemiddeld cijfer in het eerste jaar*

Wanneer het Naylor-Shine model (1965) wordt gebruikt om de toename in het gemiddelde cijfer te bepalen bij predictieve validiteit van  $r = .5$ , en een milde selectieratio van .80, is het resultaat een gemiddeld cijfer van  $M_{nieuw} = 6,8$ . Om dit te bereiken zou iedereen met een score lager dan 26 op de selectietoets afgewezen moeten worden. Wanneer deze milde selectieratio zou worden toegepast, zou het gemiddelde cijfer behaald in het eerste jaar gemiddeld met 0,2 punten stijgen. Het oorspronkelijke gemiddelde was immers  $M = 6,6$ .

Wanneer uitgegaan wordt van een strenge selectie met een selectieratio van .30, is het resultaat een gemiddeld cijfer in het eerste jaar van  $M_{nieuw} = 7,4$ . Om dit te bereiken zou iedereen met een score lager dan 33 op de selectietoets afgewezen moeten worden. Wanneer de procedure met deze strenge selectieratio zou worden toegepast, zou het gemiddelde cijfer behaald in het eerste jaar gemiddeld met 0,7 punten stijgen ( $7,3 - 6,6 = 0,7$ ). Een strenge selectie leidt tot zichtbaar succesvollere studenten, terwijl een milde selectie een kleine verbetering oplevert. Wel moet opgemerkt worden dat een strenge selectie tot meer ten onrechte afgewezen studenten zal leiden, wanneer een dichotomie tussen geschikt en ongeschikt op basis van het gemiddelde cijfer gemaakt wordt.

### **Predictieve validiteit en contextfactoren in het hoger onderwijs**

Het effect van selectie op de rendementen en de kwaliteit van studenten is afhankelijk van de predictieve validiteit van een selectieprocedure, de toevalskans en de selectieratio. Deze drie begrippen worden hieronder besproken in de context van het Nederlandse hoger onderwijs.

#### *Predictieve validiteit*

In de tabellen die bij de verschillende utiliteitsmodellen horen wordt de predictieve validiteit weergegeven in een range van nul tot één. Niet al deze waarden zijn even realistisch. De predictieve validiteit van methoden en instrumenten gebruikt in voorspellen van prestaties in het onderwijs zal nooit gelijk zijn aan 1, of daar erg dicht in de buurt komen. Afhankelijk van het criterium (succes na 1 jaar, succes na meerdere jaren studie, uitval, cijfers etc.) zal de predictieve validiteit van een optimale selectieprocedure vaak liggen tussen  $r = .30$  en  $r = .60$ .

Criteria die zich verder in de tijd bevinden zullen over het algemeen minder goed te voorspellen zijn. De onderste helft van de tabellen kunnen in deze context als onrealistisch beschouwd worden en kunnen buiten beschouwing worden gelaten.

### *Toevalskans*

Ook de toevalskans hangt sterk af van het criterium dat gebruikt wordt en kan sterk variëren tussen verschillende opleidingen. De toevalskans voor het criterium niet stoppen met de studie in het eerste jaar kan redelijk groot zijn, terwijl de toevalskans voor het behalen van een bachelor in drie jaar een stuk kleiner is (gemiddeld 33% voor studenten in het WO met startjaar 2010; Centraal Bureau voor de Statistiek, 2014)). Voor criteria met een lage toevalskans kan een veel sterkere groei in rendement verwacht worden door het invoeren van een selectieprocedure.

### *Selectieratio*

Een blik op de tabellen van de utiliteitsmodellen laat zien dat zelfs als de toevalskans laag is en de predictieve validiteit hoog, een selectieprocedure pas een behoorlijk groei in rendement zal opleveren als ook de selectieratio niet erg hoog is. Dat wordt ook geïllustreerd door de gegevens verzameld in de Verenigde Staten door ACT (2014), waarin te zien is dat onderwijsinstellingen met een hogere mate van selectiviteit lagere uitvalpercentages en hogere rendementen bereiken. Om utiliteit van een selectieprocedure te verwezenlijken moet een aanzienlijk deel van de sollicitanten worden afgewezen. Hoewel het aantal opleidingen in het Nederlandse hoger onderwijs met een numerus fixus sterk is toegenomen (Onderwijsraad, 2014), betekent dat niet dat de toegang tot het onderwijs ook selectiever is geworden. Een eenvoudige analyse op basis van gegevens uit het jaarverslag numerus fixus-opleidingen (Dienst Uitvoering Onderwijs, 2014) laat zien dat een groot deel van de opleidingen met een numerus fixus helemaal geen studenten selecteert (of uitloot, voor de opleidingen die nog geen (volledige) decentrale selectie uitvoeren).

De selectieratio voor iedere opleiding is berekend door het maximale aantal toelaatbare studenten (numerus fixus) te delen door het aantal gegadigden. In het wetenschappelijk onderwijs (WO) is er een onderscheid tussen opleidingen met een opleidingsfixus en opleidingen met een instellingsfixus. Voor opleidingen met een opleidingsfixus wordt landelijk geselecteerd en voor opleidingen met een instellingsfixus wordt per instelling geselecteerd. Van de zeven opleidingen met een opleidingsfixus was er voor twee opleidingen (biomedische wetenschappen en psychologie) sprake van een selectieratio van 1. Alle studenten die zich aanmeldden werden toegelaten. Nog eens twee opleidingen hadden een milde selectieratio groter dan .80, maar kleiner dan 1, en drie opleidingen hadden een selectieratio kleiner dan .80. Voor de 25 opleidingen met een instellingsfixus gold dat er 19 (76%) een selectieratio van 1 hadden. Bij het merendeel van de WO opleidingen met een numerus fixus werd dus niet daadwerkelijk geselecteerd. Vijf van deze opleidingen (20%) hadden een selectieratio groter dan .80 maar kleiner dan 1. Eén opleiding (4%) had een selectieratio kleiner dan .80. Het hoger beroepsonderwijs (Hbo) kent alleen een instellingsfixus. Van de 88 Hbo-opleidingen met een numerus fixus, hadden er 62 (71%) een selectieratio van 1. Elf opleidingen (13%) hadden een milde selectieratio groter dan .80 maar kleiner dan 1 en 15 opleidingen (17%) hadden een selectieratio kleiner dan .80. Van een redelijk strenge selectie (selectieratio kleiner dan .50) is slechts sprake bij de WO opleidingen diergeneeskunde en tandheelkunde en de Hbo opleidingen mondzorgkunde aan de Hogeschool Utrecht, forensisch onderzoek aan de Hogeschool van Amsterdam en medische hulpverlening aan de Hogeschool Utrecht en de Hogeschool Rotterdam. Er wordt dus weinig daadwerkelijk geselecteerd bij opleidingen met een numerus fixus. Het merendeel van de opleidingen laat alle studenten toe die zich aanmelden. Als er wel geselecteerd wordt, wordt vaak maar een klein deel van de sollicitanten afgewezen.

## Discussie

Het is van belang om naast de predictieve validiteit van een selectieprocedure, ook de selectieratio en de toevalskans mee te nemen in utiliteitsmodellen om de toegevoegde waarde in termen van rendementen en kwaliteit van studenten te bepalen. Een punt dat vaak gemaakt wordt naar aanleiding van het model van Taylor en Russell (1939) is dat instrumenten met een relatief geringe predictieve validiteit toch een grote winst kunnen opleveren als er sprake is van een lage toevalskans en/of selectieratio. Het is dan onterecht om een instrument te verwerpen op basis van een geringe predictieve validiteit alleen. Het punt dat hier gemaakt wordt is het omgekeerde daarvan. Zelfs instrumenten met een relatief hoge predictieve validiteit zullen weinig tot geen winst opleveren, als er sprake is van een hoge toevalskans en/ of selectieratio, zoals in het Nederlandse hoger onderwijs bij veel studies die een numerus fixus toepassen het geval is. Een aanzienlijk deel van de opleidingen met een numerus fixus lijkt die namelijk niet nodig te hebben, tenminste niet met de reden dat het aantal aanmeldingen het aantal plaatsen overstijgt. Als men selectie in wil zetten als effectief instrument om de rendementen en kwaliteit van studenten te verhogen, zal men ten eerste moeten zorgen voor selectie instrumenten met een hoge predictieve validiteit, waarbij in overweging moet worden genomen of men niet in te sterke mate op iets selecteert wat ook in de opleiding nog wordt aangeleerd. Daarnaast zal de selectieratio (drastisch) naar beneden bijgesteld moeten worden. Indien dit niet gebeurt is het gebruiken van selectiemethoden in dit opzicht van weinig tot geen toegevoegde waarde. Bovendien worden er bij het gebruik van selectieprocedures altijd personen ten onrechte afgewezen. Opleidingen die een numerus fixus invoeren met als doel studenten te selecteren zouden moeten afwegen of de vaak geringe stijging in rendement en kwaliteit van studenten in ethisch opzicht opweegt tegen het ten onrechte afwijzen van studenten.

## Referenties

- Abrahams, N. M., Alf, E. F., & Wolfe, J. J. (1971). Taylor-Russell tables for dichotomous criterion variables. *Journal Of Applied Psychology*, 55, 449-457. doi:10.1037/h0031761
- ACT. (2014). *National collegiate retention and persistence to degree rates*. Opgevraagd via [http://www.act.org/research/policymakers/pdf/retain\\_2014.pdf](http://www.act.org/research/policymakers/pdf/retain_2014.pdf)
- Brogden, H. E. (1949). When testing pays off. *Personnel Psychology*, 2, 171-185.
- Cascio, W. F. (1980). Responding to the demand for accountability: A critical analysis of three utility models. *Organizational Behavior & Human Performance*, 25, 32-45. doi:10.1016/0030-5073(80)90024-0
- Centraal Bureau voor de Statistiek. (2014). *WO; studievoortgang, vooropleiding, studierichting, herkomstgroepering* [Databestand]. Opgevraagd via <http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=71199ned&D1=0-21,30-53&D2=0&D3=0&D4=0&D5=a&D6=0&D7=a&VW=T>
- Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions*. Urbana, Illinois, Verenigde Staten: University of Illinois Press.
- Dienst Uitvoering Onderwijs (2014). *Jaarverslag numerus fixus-opleidingen*. Opgevraagd via [https://duo.nl/Images/Jaarverslag-numerus-fixi-studiejaar-2014-2015\\_tcm7-50219.pdf](https://duo.nl/Images/Jaarverslag-numerus-fixi-studiejaar-2014-2015_tcm7-50219.pdf)
- Drenth, P. J. D. (1995, 30 maart ). Duijkerlezing; in Nederland is selectie onmogelijk. *NRC Handelsblad*. Opgevraagd via <http://www.nrc.nl/handelsblad/van/1995/maart/30/duijkerlezing-in-nederland-is-selectie-onmogelijk-7262098>
- Drenth, P. J. D. (1998). Selectie voor en in de studie geneeskunde. *Bulletin Medisch Onderwijs*, 17, 97-107.
- Heijden, P. G. M. van der., Hessen, D.J., & Wubbels, T. (2012). Studiesucces of -falen van eerstejaarsstudenten voorspellen: een nieuwe aanpak. *Tijdschrift voor Hoger Onderwijs*, 4, 233-244.
- Lawshe, C. H., Bolda, R. A., Brune, R. L., & Auclair, G. (1958). Expectancy charts: II. Their theoretical development. *Personnel Psychology*, 11, 545-559. doi:10.1111/j.1744-6570.1958.tb00040.x
- McLellan, R. A. (1996). Theoretical expectancy calculator [online utility calculator]. Opgevraagd via <http://www.hr-software.net/cgi/TheoreticalExpectancy.cgi>
- Meijer, R. R. & Niessen, A. S. M. (2015). A trial studying approach to predict college achievement. *Frontiers in Psychology*, 6:887. doi: 10.3389/fpsyg.2015.00887

- Mellenbergh, G. J. (1993). Beslissen met tests en studietoetsen. In: Koele, P. & van der Pligt, J. (1993). *Beslissen en beoordelen*. Amsterdam: Boom, p. 96-124.
- Mellenbergh, G. J. (1995). Selectie aan de poort. *Spiegeloo*, 8-9.
- Ministerie van Onderwijs, Cultuur en Wetenschap (29 augustus 2014). *Informatie over de afschaffing van loting bij numerusfixusopleidingen*. Kamerbrief. Opgevraagd via <http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2014/08/30/kamerbrief-met-informatie-over-de-afschaffing-van-loting-bij-numerusfixusopleidingen.html>
- Naylor, J. C. & Shine, L. C. (1965). A table for determining the increase in mean criterion score obtained by using a selection device. *Journal of Industrial Psychology*, 3, 33-42.
- Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2014). *Predicting success in higher education using proximal predictors*. Artikel ingediend ter publicatie.
- Onderwijsraad. (2014). *Overgangen in het onderwijs*. Opgevraagd via <http://www.onderwijsraad.nl/publicaties/2014/overgangen-in-het-onderwijs/volledig/item7086>
- Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological Bulletin*, 138, 353–387. doi:10.1037/a0026838
- Sackett, P. R., & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology*, 85(1), 112-118. doi:10.1037/0021-9010.85.1.112
- Schripsema, N. R., van Trigt, A. M, Borleffs, J. C. C., & Cohen-Schotanus, J. (2014). Selection and study performance: comparing three admission processes within one medical school. *Medical Education*, 48, 1201-2010. doi: 10.1111/medu.12537
- Smith, M. (1948). Cautions concerning the use of the Taylor-Russell tables in employee selection. *Journal of Applied Psychology*, 32(6), 595-600. doi:10.1037/h0058507
- Taylor, H. C., & Russell, J. T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection: discussion and tables. *Journal of Applied Psychology*, 23(5), 565-578. doi:10.1037/h0057079
- Urlings-Strop, L. C., Stijnen, T., Themmen, A. N., & Splinter, T. W. (2009). Selection of medical students: A controlled experiment. *Medical Education*, 43, 175-183. doi:10.1111/j.1365-2923.2008.03267.x

Visser, K., van der Maas, H., Engels-Freeke, M. & Vorst, H. (2012). Het effect op studiesucces van decentrale selectie middels proefstuderen aan de poort. *Tijdschrift voor Hoger Onderwijs* 30, 161-173.

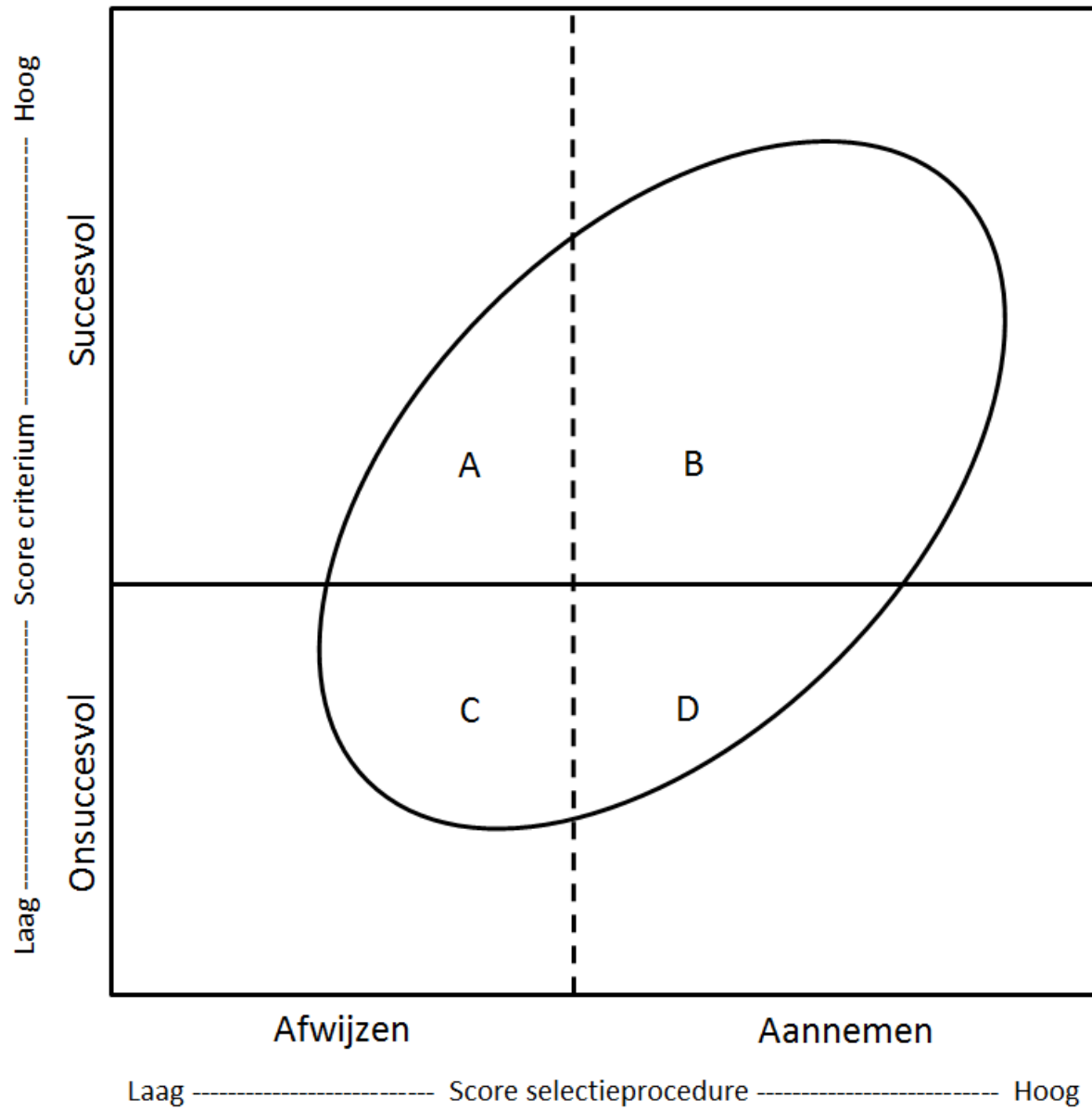
Wet kwaliteit in verscheidenheid (2013). *Memorie van Toelichting*. Opgevraagd via [https://www.eerstekamer.nl/behandeling/20130118/memorie\\_van\\_toelichting\\_2/document3/f=/vj6ifr0xnkyo.pdf](https://www.eerstekamer.nl/behandeling/20130118/memorie_van_toelichting_2/document3/f=/vj6ifr0xnkyo.pdf)

$$(1) \quad \textit{Succesratio} = \frac{B}{B+D}$$

$$(2) \quad \textit{Toevalskans} = \frac{A+B}{A+B+C+D}$$

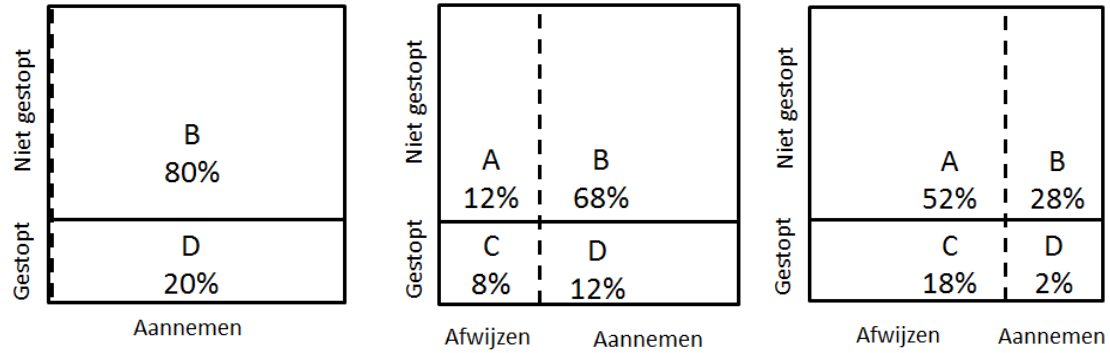
$$(3) \quad \textit{Selectieratio} = \frac{B+D}{A+B+C+D}$$

*Figuur 1.* Vergelijkingen behorend bij het Taylor-Russell model (1939)



*Figuur 2.* Schematische weergave van het effect van cutoff scores van de selectieprocedure en het criterium (naar Taylor en Russell, 1939)





*Figuur 3.* Percentage personen in iedere categorie bij (a) geen selectie, (b) milde selectie (selectieratio = .80) en (c) strenge selectie (selectieratio = .30).